Computer Science

SOFTWARE FOR READING METABOLIC RECONSTRUCTION
INTO A DATABASE INDEPENDENT OBJECT MODEL

John Ayettey, Soon-Ok Park*, Governors State University, University Park, IL 60466
Maurice McDowell, Chicago State University, Chicago, IL 60628
Edward Frank*, Michael Papka*
* Mathematics and Computer Science Division,
Argonne National Laboratory, Argonne, IL 60439

The field of Systems Biology builds models of biological processes based upon experimental biological data. It is anticipated that these models will grow in size and complexity as we understand more about the dynamical aspects of these systems. These models presently use data that are spread over multiple databases. These databases represent similar data in dissimilar ways. Also, the structure of the databases change over time. Dissimilar database structures makes it difficult to construct models and the evolution of the databases over time result in wasted effort maintaining applications that use the data.

A database-independent object model of the biological information can be used to insulate the application from these forces. In this project, we constructed software to read data from two dissimilar Oracle databases into a common object model.

We first investigated technological choices for the project, comparing ODBC based database access with Python DBI based database access. We found the Python DBI to be a much simpler, but less capable, program interface but, since it met our needs, we chose Python and Python-DBI as the technological underpinnings for the project. We then implemented the object model in Python, wrote so-called "factories" to read data from the two databases and convert the data into instantiations of the objects in the object model. We wrote test programs and, at the end of the project, connected the software to an independently produced application that was developed in terms of the object model. This application, a Python based Model Editor, will be improved in future work to allow construction of models from data in multiple databases.